

DOCUMENT RESUME

ED 409 311

SP 037 444

AUTHOR De Corte, Erik; And Others
TITLE Learning Bad and Good Things from Instruction: A European Perspective.
PUB DATE Aug 96
NOTE 37p.; Paper presented at the Annual Meeting of the American Psychological Association (Toronto, Ontario, Canada, August 9-13, 1996).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE EDRS Price MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Arithmetic; Elementary School Mathematics; Elementary School Students; Foreign Countries; Grade 5; Grade 6; Higher Education; Intermediate Grades; *Mathematical Models; Mathematics Skills; *Problem Solving; *Student Attitudes; Student Teachers; *Theory Practice Relationship; *Thinking Skills; *Word Problems (Mathematics)
IDENTIFIERS Belgium; Japan; Preservice Teachers

ABSTRACT

This paper reports on five related studies on upper elementary school students' beliefs about the role of real-world knowledge in solving mathematical word problems, and on possibilities of overcoming their unrealistic mathematical modeling. Findings from the first study of 75 Belgian fifth-grade students showed that as a consequence of current classroom practice in teaching arithmetic word problem solving, pupils have built up a tendency to solve word problems in a stereotyped way, i.e., they applied one of the basic arithmetic operations (or a combination) with the two given numbers in the problem, without any consideration of the possible problematic modeling assumptions underlying their proposed solution. The results of the second and third studies (64 Belgian and 91 Japanese fifth-grade students) confirmed this tendency, and indicated that it is culturally independent. The fourth study involving 332 Flemish preservice teachers revealed a similar tendency as reflected by these future teachers' preferred spontaneous solutions to word problems and their evaluations of different types of pupil answers. In the fifth study, 53 Belgian sixth-grade students were immersed in a new classroom culture in which word problems were treated as exercises in realistic mathematical modeling. The findings from the fifth study supported the hypothesis that, with specific instructional techniques and sets of challenging problem situations, it is possible to change pupils' conceptions about and their disposition toward the activation of real-world knowledge in mathematical modeling of word problems. (Contains 27 references.) (ND)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

LEARNING BAD AND GOOD THINGS FROM INSTRUCTION: A EUROPEAN PERSPECTIVE

Erik DE CORTE, Lieven VERSCHAFFEL, Sabien LASURE, Inge BORGHART

Center for Instructional Psychology and Technology (CIP&T)

University of Leuven, Belgium

&

Hajime YOSHIDA

Department of Psychology, Faculty of Education

Miyazaki University, Japan

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☐ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.

* Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

August 1996

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL
IN OTHER THAN PAPER COPY
HAS BEEN GRANTED BY

E. De Corte

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

Invited address for Division 15: Educational Psychology of the American Psychological Association (APA), organized at the Annual Convention of APA held in Toronto, Canada, August 9-13, 1996

Address for correspondence: Erik De Corte, Center for Instructional Psychology and Technology, University of Leuven, Vesaliusstraat 2, B-3000 Leuven, Belgium. Tel: int.-32-16-32.62.48 - Fax: int-32-16-32.62.74 - E-mail: Erik.DeCorte@ped.kuleuven.ac.be

5637444

LEARNING BAD AND GOOD THINGS FROM INSTRUCTION: A EUROPEAN PERSPECTIVE

Introduction

Analyses of expert performance in a large variety of disciplines has shown that becoming skilled in a domain requires the integrated acquisition of the following four categories of aptitudes: a well-organized and flexibly accessible domain-specific knowledge base, heuristic methods for problem analysis, metacognitive knowledge and skills, and affective components such as positive beliefs, attitudes, motives, and emotions relating to subject-matter fields. In addition, to overcome the well-known phenomenon of inert knowledge one should acquire a disposition toward skilled learning, thinking, and problem solving, which refers especially to the sensitivity to situations in which it is relevant and useful to apply one's knowledge and skills, on the one hand, and to the inclination to do so whenever appropriate, on the other (for a brief review of the literature, see De Corte, 1995b, 1996).

There is nowadays ample evidence showing that students in today's schools do often not master sufficiently the categories of knowledge and cognitive skills mentioned above to approach new learning tasks and problem situations in an efficient and successful way (see e.g., De Corte, 1995a). For instance, students often acquire only deficient, superficial, and rote knowledge of basic concepts, and in many domains they are even affected with misconceptions (e.g., "multiplication makes bigger") and defective skills (e.g., buggy algorithms); they also do not master a variety of usable heuristic methods, and in their learning and thinking processes there are very few traces of metacognitive and self-regulating activities. Interestingly, those shortcomings and deficiencies can be largely attributed to current teaching practices; thus, in the cognitive domain, students seem to learn bad things from instruction.

But what about the more affective aspects that are involved in learning, understanding, and thinking? As compared to the cognitive components, it is rather recently that researchers have begun to study more thoroughly the affective aspects that are involved in expert performance in a domain (see e.g., the book edited by McLeod & Adams, 1989, relating to the domain of mathematics). This work has meanwhile already documented that apparently also in the affective domain, students learn bad things from instruction. For instance, Schoenfeld (1988) has shown that in secondary school mathematics classes, where teaching is generally considered to be of good quality, students nevertheless acquire strange and uncorrect beliefs about mathematics problem solving, such as "being able to solve a math problem is mainly a matter of

being lucky", and "solving an assigned math problem should not take more than five minutes. Over the past years researchers in several European countries (Belgium, Sweden, Switzerland, and Northern Ireland) have studied rather intensively one aspect of the belief system of mainly primary school children, namely their beliefs about the role of real-world knowledge in mathematical modeling of school word problems. This investigations have so far yielded substantial empirical evidence that students at different levels of the school system - seemingly as a consequence of current teaching practices in the mathematics classroom - develop the belief that real-world knowledge and realistic considerations should be neglected when solving arithmetic word problems. For instance, when asked how much time it will take someone to run 1 km given that his best time to run 100 m is 17 seconds, the majority of the children at the end of the primary school answer without any hesitation: 170 seconds. The most spectacular illustration of children's failure to make proper use of real-world knowledge and sense-making activities during problem solving has been reported already in the 1980s by French and German researcher. They administered absurd, unsolvable tasks, like "There are 26 sheep and 10 goats on a ship. How old is the captain?", and found that many primary school children came up with an answer resulting from making computations with the given numbers; the children seemed not to be aware of the absurdity of the task nor of the meaninglessness of their solution (for a review of these studies see Selter, 1994).

In some way these observations confront us with a kind of paradox. Indeed, a major argument for the inclusion of verbal problems in the mathematics curriculum has always been to develop in pupils the skill of using and applying their mathematics knowledge to model and solve tasks and problems encountered in real life (see e.g., Burkhardt, 1994). However, it looks as if current practice of teaching word problems in schools does not at all foster in pupils a disposition to realistic modeling and problem solving. To the contrary, instead of inviting children to rely on, and employ their real-world knowledge and experience during problem solving, school word problems are perceived as artificial tasks that are unrelated to the realities of life outside the school (Freudenthal, 1991; Greer, 1993; Reusser, 1988; Treffers, 1987; Verschaffel & De Corte, in press).

In this presentation I will focus on a series of recent studies carried out in our Center in Leuven that aimed at a more detailed description and understanding of students' beliefs about the role of real-world knowledge during solving arithmetic word problems, but also at exploring possibilities to overcome their unrealistic mathematical modeling. First, I will describe two studies that provide strong evidence for the position that as a result of traditional math education children develop a narrow, but deep-rooted and resistant view of mathematical modeling. In this view modeling is reduced to selecting the correct formal arithmetic operation with the given

numbers, without seriously taking into account their common-sense world knowledge and realistic considerations about the problem context. Interestingly, in a third study the same phenomenon was observed in a group of Japanese children. In a fourth study pre-service teachers were confronted with the same problems as the children from the previous study. Taking into account our better understanding of the phenomenon of unrealistic modeling derived from the preceding investigations, we carried out an intervention study to test the hypothesis that it is possible to change through instruction children's (mis-)beliefs about the role of real-world knowledge in mathematical modeling and problem solving. The presentation will end with some conclusions and discussion.

The phenomenon of unrealistic modeling in fifth graders

Study 1: Ascertaining the phenomenon

Method

The subjects were 75 pupils (10-11 year-old boys and girls) from three fifth-grade classes. They were taught word problem solving using traditional rather than authentic problem situations, and realistic mathematical modeling was not systematically addressed in teaching (see also De Corte & Verschaffel, 1989)

A paper-and-pencil test was constructed consisting of ten matched pairs of items. Each pair consisted of

- a standard item (S-item) that can be solved unambiguously by applying the most obvious arithmetic operation(s) with the given numbers (e.g., "Steve has bought 5 planks of 2 meters each. How many planks of 1 meter can he saw out of these planks?"), and
- a parallel problematic item (P-item) for which the appropriate mathematical model is less obvious and indisputable, at least if one seriously takes into account the realities of the context evoked by the problem statement (e.g., "Steve has bought 4 planks of 2.5 meters each. How many planks of 1 meter can he saw out of these planks?").

The ten P-items are listed in Table 1.

Insert Table 1 about here

The ten pairs of problems were administered by the class teacher on the same day in two series, each containing the P-variant of five problem pairs and the S-variant of the five other pairs. With respect to each problem, pupils were asked to write down not only their answer, but also how they arrived at this answer (e.g., by mentioning the calculations), and possible other comments they might have (e.g., explaining their stumbling block when they were not able to solve the problem, supplementing their numerical answer with some comments, criticizing the problem statement, etc.).

Analysis

Children's reactions to the problems were analyzed in two ways for evidence of the activation and use of real-world knowledge and realistic considerations about the problem context:

- (1) by distinguishing in their answers between realistic answers and non-realistic ones;
- (2) by distinguishing in their computations and additional comments between realistic comments and no realistic comments.

When a child gave an answer to the problem that was scored as realistic or produced a non-realistic answer that was accompanied by a realistic comment, his(her) overall reaction to that particular problem was scored as a "realistic reaction" (RR). Take, for example, the "planks" item mentioned above. A RR score was not only given to a child who produced the realistic answer "8 planks", but also to a child who responded with "10 planks" but who added the comment that "Steve would have a hard time putting together the remaining pieces of 0.5 meter". The code NR ("non-realistic reaction") was given for children who answered a problem in a non-realistic manner, and did not give any further realistic comment. (For examples of realistic and non-realistic reactions to the ten P-items, see Verschaffel, De Corte, & Lasure, 1994).

Hypotheses

The overall hypothesis of the first study was that due to their extensive experience with an impoverished diet of standard word problems, and to the lack of systematic attention at the mathematical modeling perspective in their mathematics lessons, pupils will demonstrate a strong tendency to exclude real-world knowledge and context-bound considerations from their problem-solving endeavours of the P-items, and - consequently - will solve them as if they were not at all problematic.

Results

Table 2 gives the number of pupils who reacted in a realistic (RR) and in a non-realistic (NR) way for each of the ten P-items from Table 1.

Insert Table 2 about here

The data in Table 2 strongly support the hypothesis. As predicted, the pupils demonstrated a very strong overall tendency to exclude real-world knowledge and realistic considerations when confronted with the problematic versions of the problems. In total, only 128 out of the 750 reactions to the P-items (= 17 %) could be considered as realistic (RR), either because the pupil wrote a realistic answer or made an additional realistic comment. Very similar results have been reported by Greer (1993) and Reusser (in press) for North-Irish and Swiss children, respectively.

Table 2 also shows that for two out of the ten P-items a considerable number of realistic answers or comments were observed: the "buses" item (P4) and the "balloons" item (P7). The question arises why these two problems elicited considerably more RR than the other P-items. A plausible explanation is that in P4 and P7 the modeling difficulty is restricted to the last phase of the solution process, wherein the pupil has to make sense of the result of the arithmetic operation, namely a quotient with a remainder (Silver, Shapiro & Deutsch, 1993). In all other P-items the underlying realistic modeling difficulty requires context-based adaptations at the beginning of the solution process, namely the construction of the mathematical model, rather than at the end of it (Verschaffel et al., 1994).

Finally, to have an indication of the interindividual differences in the disposition toward (non-)realistic modeling, we counted the total number of RRs on the ten P-items for each pupil separately. Fifty-nine out of the 75 pupils (= 78 %) reacted in a realistic way to less than three of the ten P-items. Almost all RRs of these pupils were found on the "buses" and the "balloons" problem. Only 16 pupils (= 22 %) provided RRs on at least three P-items, and only two of them did so on more than half of the ten P-items (see Verschaffel et al., 1994, for more details).

Discussion

While Study 1 provides empirical evidence for pupils' strong tendency to exclude real-world knowledge and realistic considerations from their understanding and solving of word problems,

its findings need to be put into perspective because of an important methodological limitation of the data-gathering technique used, namely a collective paper-and-pencil test. Indeed, one could argue that during their (private) solutions of the P-items some pupils may have activated real-world knowledge which was not reflected in their written answers, simply because they finally decided to react in a "conformist" rather than a "realist" way in line with their beliefs and conceptions about "the prevailing rules and conventions of the game of school arithmetic word problems" (De Corte & Verschaffel, 1985). This may have led to a significant underestimation of the number of realistic considerations in this investigation. Therefore, we set up a second study in which we assessed whether two rather simple forms of scaffolding during an individual interview are sufficient to transform children's non-realistic solutions into realistic ones.

Study 2: Do weak forms of scaffolding help?

Method

Study 2 consisted of two stages. In the first stage, seven matched pairs of word problems used in Study 1 were collectively administered to a group of 64 fifth-graders from three classes. Based on the results on this paper-and-pencil test, the five most "realistic" and the five most "non-realistic" problem solvers from each of the three fifth-grade classes were selected to participate in the second stage of the investigation. During this second stage, which took place one or two days later, these 15 realistic and 15 non-realistic problem solvers were individually administered the same seven problem pairs once again.

To assess the strength of children's tendency towards non-realistic mathematical modeling as well as the awareness of the beliefs and strategies underlying their non-realistic solutions, the following interviewing procedure was followed with respect to each P-item solved with a NR on the paper-and-pencil test.

First, the pupil was asked to read aloud the problem followed by his own NR written down on the answer sheet (e.g., " $4 \times 2,5 = 10$ planks" for the "planks" item).

Then the pupil was confronted with the written notes of a fictitious classmate who had responded in a realistic manner. For instance, with respect to the "planks" item the interviewer said: "As you can see on this sheet, one of your classmates responded: $4 \times 2 = 8$ planks. What is the best answer? Why?". If the pupil stuck to the initial NR (i.e., "10 planks") after this first and weak form of scaffolding, a second and stronger scaffold in the direction of realistic modeling was provided by stimulating and helping the pupil to realize the concrete problem situation. For example, with respect to the "planks" item the interviewer said: "Can you draw the planks? Can

you also draw what happened with these planks according to the problem statement? Can you see on your drawing how many planks of 1 meter Steve can saw out of these 4 planks?".

Analysis

For all 30 pupils and for all seven P-items the following kinds of data were used for analysis:

- (1) the initial reaction to the problem during the paper-and-pencil test as (RR or NR);
- (2) the reaction to the confrontation with the RR of a fictitious classmate (= weak scaffold);
- (3) eventually, the reaction to the second and strongest form of scaffolding.

While the pupils' total number of RRs to the seven P-items of the paper-and-pencil test provide a measure of their actual level of realistic modeling - i.e. the level at which they can perform independently -, their sensitivity to scaffolds towards realistic modeling of the P-items that were initially answered with a NR reflect their zone of proximal development (Vygotsky, 1978) or their learning potential (Hamers, Sijtsma, & Ruijsenaars, 1993), i.e., those behaviors which they cannot yet perform independently but can with help from others. To measure this zone of proximal development or learning potential, a score of 2, 1 or 0 was given for each P-item not yet solved in a realistic manner during the paper-and-pencil test:

- 2 = the pupil replaced the NR during the paper-and-pencil test by a RR in response to the first, weak form of scaffolding;
- 1 = the pupil produced a RR only after receiving the second, stronger scaffold;
- 0 = the pupil's NR remained unchanged even after the second and strongest scaffold.

The mean of these scores (= their sum divided by the number of P-items not yet answered with a RR during the paper-and-pencil test) was considered as a measure of a pupil's learning potential, i.e. a score indicating how much a pupil gained from increasingly stronger forms of scaffolds toward realistic modeling on those P-items he could not answer in a realistic manner himself.

Hypotheses

First, it was hypothesized that the overall number of RRs generated on the seven P-items of the paper-and-pencil test would be alarmingly low, and, thus would not differ significantly from the percentage found in Study 1, which was 17 %.

Second, a positive effect of the scaffolds on the number of RRs was anticipated. Therefore, we predicted that the percentage of RRs of the 30 pupils at the end of the individual interview would be significantly higher than on the paper-and-pencil test. However, we also expected that this overall percentage of RRs at the end of the individual interviews would still be quite low. This prediction was based on the assumption that pupils' tendency towards routine-based and non-realistic modeling would be so strong, that the confrontation with the RR (= scaffold 1) and even the subsequent hint toward the realistic reasoning underlying this RR (= scaffold 2) would often be insufficient to make them change their initial NR into a RR.

Third, we hypothesized that the so-called realistic problem solvers of the P-items from the paper-and-pencil test would benefit more from the two forms of scaffolding than the pupils with little or no RRs on that test (= the so-called non-realistic problem solvers). Their greater initial tendency towards realistic modeling would make them also more sensitive to the two scaffolds provided during the individual interview. Therefore, we predicted a significantly greater mean learning potential score for the 15 realistic than for the 15 non-realistic problem solvers on the P-items not yet solved with a RR during the paper-and-pencil test.

Results

In line with our first hypothesis and with the findings of Study 1, the results on the collective test revealed a very strong tendency among the pupils to exclude real-world knowledge and context-based considerations from their solutions of school arithmetic word problems: only 16 % of all the reactions of the 64 pupils to the seven P-items of the paper-and-pencil test were classified as realistic. This percentage is almost exactly the same as the overall percentage in Study 1 (i.e. 17 %).

As mentioned before, from each of the three participating classes the five pupils with the highest number of RRs and the five pupils with the lowest number of RRs were selected to participate in the second part of the study. The percentage of RRs of those 15 realistic and 15 non-realistic problem solvers was 39% and 8%, respectively. A one-tailed t-test showed that the difference between the two groups was significant, $t(28)=8.58$, $p < .01$.

In accordance with the second hypothesis, a significant effect of the two forms of scaffolding was found. Altogether, the two scaffolds resulted in an increase in the cumulative percentage of RRs from 23 % (i.e. 49 RRs out of a total 210 responses) during the paper-and-pencil test to 57 % (i.e. 120 RRs out of a total of 210 responses) at the end of the individual interviews (see Table 3). A one-tailed t-test revealed that this increase was significant, $t(29)=12.97$, $p < .001$. Additional one-tailed t-tests revealed that both forms of scaffolding contributed equally to this

increase: the first, weak scaffold resulted in an increase of the cumulative percentage of RRs from 23 % (i.e. 49 RRs out of 210 responses) to 40 % (i.e. 83 RRs out of 210 responses), $t(29)=5.61$, $p < .001$, and the second and stronger scaffold produced an additional significant increase from 40 % (i.e. 83 RRs out of 210 responses) to 57 % RRs (i.e. 120 RRs out of 210 responses), $t(29)=6.95$, $p < .001$.

Insert Table 3 about here

Although the scaffolds led to a significant increase in the number of RRs, the cumulative percentage of RRs at the end of the individual interviews was still alarmingly low. Indeed, in 43 % of all cases (i.e. 90 out of a total of 210) the pupils still reacted in an unrealistic way even after the second and strongest scaffold towards realistic modeling. Moreover, a qualitative analysis of the pupils' reactions to the two forms of scaffolding indicated that pupils frequently stuck to their NR after the first and the second form of scaffolding, because they did not seem to grasp the line of reasoning underlying the RR proposed by the interviewer (= scaffold 1), even not after the interviewer's suggestion to realize the problem situation (= scaffold 2).

Finally, the results confirmed also the hypothesis that the realistic problem solvers would benefit more from the two scaffolds than the non-realistic ones. This was tested by comparing the mean learning potential scores (maximum= 2; minimum= 0) of the 15 realistic and the 15 non-realistic problem solvers. The mean learning potential score was .88 and .52 for the realistic and the non-realistic group, respectively. A one-tailed t-test revealed that this difference was significant, $t(28)=2.55$, $p < .01$.

Thus, while the results of the second study strongly confirm elementary school pupils' disposition towards non-realistic modeling of school arithmetic word problems, they also show that this tendency is very strong and resistant to change (see also Reusser, in press). In addition, the findings provide evidence for the existence of interindividual differences in this disposition, by showing that pupils of different levels of actual performance with respect to realistic modeling, differ also in terms of their zone of proximal development, in the sense that they are differently sensitive to scaffolds towards realistic modeling on problems they could not solve properly themselves initially.

Study 3: Comparing Belgian and Japanese children

On the basis of the overall picture that derives from the large-scale cross-cultural research on mathematics learning and teaching over the past decades (see e.g., Robitaille & Travers, 1992; Stevenson & Stigler, 1992; for a summary see De Corte, Greer, & Verschaffel, 1996), one might suspect that Japanese pupils will perform significantly better on the P-items used in the previous studies as compared to European children. But, the findings of a few more recent, small-scale, but at the same time more focused studies (e.g., Cai & Silver, 1995) indicate that the differences in performance on this type of problems may be rather small, or even nonexistent.

Against this background the present investigation was designed to examine Japanese children's performance on a test of realistic mathematical modeling in comparison with the Belgian pupils tested in Study 1. A second purpose was to investigate the effect on Japanese pupils' test results of a similar kind of weak scaffold as used in Study 2 with Belgian children, namely a simple but very explicit hint given at the beginning of the test in order to induce a more mindful and more realistic approach to the problems. This hint consisted in warning the pupils that the test would contain some items that are difficult or impossible to solve, and inviting them to comment on those items explaining why they are unsolvable.

Method

The subjects were 91 fifth graders from three classes in the same Japanese school located in a suburb of a middle-size city. They were randomly assigned to two conditions: the 45 children in condition 1 received neutral instructions as in Study 1; the 46 pupils in condition 2 were warned of the presence of problematic items in the test.

The paper-and-pencil test that was administered to the pupils consisted of literal translations in Japanese of the ten pairs of items used in Study 1, except that names of children were replaced by typical Japanese names. As an illustration Figure 1 shows the Japanese translation of item P2 from Table 1.

Insert Figure 1 here

The procedure for administering the test in condition 1 was the same as in Study 1. In condition 2 the students were given in addition the "warning"-instruction which was written on top of the

test booklet. Children's reactions to the problems were also scored in the same two categories as in Study 1: realistic reactions (RR) and non-realistic reactions (NR).

Results

To assess the tendency to exclude real-world knowledge during problem solving, the number and percentage of RR were computed for the ten P-items separately, and over the ten P-items in conditions 1 and 2. These data were compared with those of the Belgian children from Study 1. The results are shown in Table 4.

Insert Table 4 here

A X^2 -analysis indicated that there were no statistically significant differences between the total percentages of RRs over the ten P-items in the three groups. With respect to the separate items a significant difference was found for the "runner"-item (P5), $X^2=10.7$, $df=2$, $p<.01$, and for the "balloons"-item (P7), $X^2=6.3$, $df=2$, $p<.05$. In both cases the percentage of RRs in condition 2 in Japan was significantly higher than in the other two groups, which did not differ from each other.

Thus, with respect to the first objective of this study, the comparison of the performance of the Japanese fifth graders in condition 1 with the Belgian children from Study 1 tested under the same condition, revealed that the tendency to neglect real-world knowledge and realistic considerations while solving school word problems was equally strong in both cultures.

With regard to the second purpose of this investigation, the results indicated that the "warning"-instruction used in condition 2 to induce more realistic modeling in view of increasing the number of RRs on the P-items, was only marginally successful, and remained restricted to only two out of the ten items. This finding is also in line with the weak effects of the two forms of scaffolding observed in Study 2, and similar results in Swiss children reported by Reusser (in press).

Discussion

The results of the three studies reported above provide strong support for the hypothesis that upper elementary school pupils have developed - independent of cultural and curricular differences - a strong belief that real-world and everyday knowledge is irrelevant when solving

school word problems. Indeed, these studies do not only show that children's answers to the P-items from the test that was administered, only rarely reflect adjustments for contextual constraints based on the activation and use of real-world knowledge. Moreover, Studies 2 and 3 evidence that children's belief about the irrelevance of real-world knowledge in problem solving is very strong and resistant to change; this derives from the observation, in Belgium as well as in Japan, that simple but explicit forms of scaffolding aimed at inducing more realistic modeling, have very poor effects. The qualitative data collected in Study 2 suggest that children's persistence to give non-realistic answers is due to their routine-based approach to the problems in which there is hardly any space for the activation and use of real-world knowledge and context-based considerations about problem situations.

The findings of the previous investigations raise the question concerning the causes of the phenomenon of unrealistic modeling in upper elementary school pupils. In this respect, many researchers attribute the development of the strong belief about the irrelevance of real-world knowledge in solving school word problems to the following major and related characteristics of the typical instructional environment in today's mathematics classroom: (1) the impoverished and stereotyped diet of standard word problems which can always be modeled and solved through the straightforward application of one or more arithmetic operations on the given numbers; (2) the nature of the instructional activities relating to word problems, especially the premature imposing of the formal arithmetic approach toward word problems by requiring that pupils must identify and execute the correct arithmetic operation to solve a word problem, and - more generally - the absence of systematic attention in teaching to the modeling perspective as one of the building blocks of a genuine mathematical disposition (De Corte & Verschaffel, 1985; Freudenthal, 1991; Greer, 1993; Kilpatrick, 1987; Reusser, 1988; Säljö, 1991; Verschaffel & De Corte, in press). However, until now the empirical evidence to support the claim that children's unrealistic modeling is due to these features of the instructional environment is rather scarce and anecdotal, and, therefore, not very compelling. The two remaining studies to be presented attempt to contribute to fill this gap in two ways: (1) by focusing on a major component of the instructional environment, namely the teacher and more specifically (student-)teachers' own beliefs about the role of real-world knowledge in problem solving (Study 4); (2) by testing the hypothesis that it is possible to develop in primary school children a disposition toward more realistic mathematical modeling by immersing them into a more appropriate and powerful teaching-learning environment (Study 5).

Student teachers' beliefs about the role of real-world knowledge in mathematical modeling

In this study pre-service teachers' own beliefs about the role of real-world knowledge and realistic considerations in modeling school word problems were studied by analyzing: (1) their own spontaneous responses to a set of word problems with problematic modeling assumptions (P-items); (2) their evaluations of pupil answers to the same problems that do or do not take into account real-world knowledge.

Method

Participants were 332 pre-service teachers from three Flemish teacher training institutions. About two thirds ($n=228$) had just started their first year of teacher training, while one third ($n=104$) were in the third, and last year of their pre-service training.

In this investigation the same paper-and-pencil test as in Study 2 was used. It consisted of seven standard (S-items) and seven problematic items (P-items). This test was administered twice to all the pre-service teachers. The first time (= Test 1), the student-teachers had to answer the 14 word problems themselves; calculations and comments could be written down in a "comments box" below the "answer box". Immediately after they had finished and handed in this test, they were given Test 2, in which they were asked to score four different answers from pupils to the same 14 word problems as in Test 1 with either 1 point ("absolutely correct answer"), 0 points ("completely incorrect answer") or 1/2 point ("partly correct and partly incorrect answer"). The four response alternatives to the seven P-items in Test 2 belonged to four different categories: a non-realistic answer (NA), a realistic answer (RA), a technical error (TE) and another answer (OA) such as a wrong-operation or a give-number answer (see Figure 2).

Insert Figure 2 here

At the bottom of each problem, there was a box for writing explanations or comments.

For Test 1, an assessment of the student-teachers' spontaneous solutions to the P-items was made involving the same two major categories as in the previous studies: realistic reaction (RR) versus non-realistic reaction (NR). The analysis of the pre-service teachers' reactions to the seven P-items in Test 2 focused on the score (1, 1/2 or 0) given to the realistic answer (RA) and the non-realistic answer (NA).

Hypotheses and questions

Hypotheses and questions with respect to Test 1. The first hypothesis with respect to Test 1 was that - due to their continuing experience with an impoverished diet of standard word problems and to the lack of systematic attention at the mathematical modeling perspective in their mathematics (education) lessons - even the student-teachers would demonstrate a strong tendency to exclude real-world knowledge when confronted with the problematic versions of the problems, and, consequently would solve these problems as if they are not at all problematic. Although we expected that the results of these 18 to 21-year-old pre-service teachers would be considerably better than those of the upper elementary school and lower secondary school pupils from the studies of Greer (1993), Reusser (in press) and Verschaffel et al. (1994), we still predicted an alarmingly low percentage of RRs on the seven P-items from Test 1.

Second, we wondered whether there would be a difference between the students who had almost completed their (third year of) training as an elementary school teacher and the students who had just started their (first year of) training. But, because of the possibility of conflicting hypotheses in this respect, no specific prediction was made about the difference in the percentage of RRs between the first-year and the third-year students (see Verschaffel, De Corte, & Borghart, in press).

Hypotheses and questions with respect to Test 2. First, we hypothesized that the student-teachers' scores of the realistic (RA) and the non-realistic answer (NA) to a P-item from Test 2 would generally reflect their spontaneous solutions of the same problems during Test 1. Consequently, we anticipated that the student-teachers would frequently consider the NA for the seven P-items from Test 2 as the (perfectly) correct answer and, therefore, score it with 1, while the RA would frequently be conceived as an inappropriate response and, thus, be scored with a 0.

Second, we wondered again whether there would be a difference in the scores for the NAs and the RAs between the first-year and the third-year students.

Hypotheses and questions with respect to the relationship between Test 1 and Test 2. A final set of hypotheses and questions concerns the relationship between a student-teacher's answer to a P-item in Test 1, on the one hand, and his or her score for the NA and the RA on the same item in Test 2, on the other hand. Generally speaking, we expected a good match between the

student-teachers' reactions to each of the seven P-items in both tests. This implies the following two predictions. First, most subjects who answered a P-item from Test 1 with a NR, will give a 1-score for the NA and a 0-score for the RA for the same item in Test 2. Second, a RR on a P-item during Test 1 will typically be followed by a 0 for the NA and a 1 on the RA for the same item during Test 2.

But, we also anticipated a number of learning and thinking processes in the student-teachers that could lead to a mismatch between a student-teachers' own response to a P-item during Test 1, and the way in which (s)he scored the realistic and the non-realistic response alternative for the same item in Test 2. A first possible source of such mismatches is the scaffolding effect of the confrontation with the realistic response alternative in Test 2. A second possible source of mismatches derives from student-teachers' pedagogical conceptions and beliefs about the importance that should be attributed to real-world knowledge in teaching school arithmetic word problem solving. For example, some subjects who had themselves demonstrated awareness of the problematic nature of a P-item during Test 1, may nevertheless give scores for the RA and for the NA reflecting little or no appreciation for pupil answers based on realistic context-based considerations, because they believe that the activation of such considerations should not be stimulated but discouraged in elementary school mathematics.

Results

Test 1. As expected, the student-teachers demonstrated a strong overall tendency to exclude real-world knowledge and realistic considerations when confronted with the problematic word problems. Indeed, only 48 % of all reactions to the seven P-items from Test 1 could be considered as realistic (RRs). This percentage is considerably higher than in previous investigations with upper elementary (see Studies 1 and 2) and lower secondary school pupils (Greer, 1993), where overall percentages of RRs between 15 % and 20 % were observed. Nevertheless, it remains disappointingly low, as it implies that in more than half of the cases, the student-teachers solved the P-items from Test 1 in a stereotyped, uncritical way, without any consideration for the realities of the context involved in the problem.

There was a significant difference in the overall number of RRs between the first-year and the third-year student-teachers in favor of the latter group (t-test, two-tailed, $t = 3.40$, $p < .01$). However, the overall percentage of RRs remained low in both groups, namely 45 % and 54 % for the first-year and the third-year students, respectively. Interestingly, the size of the difference between both years was dissimilar for the three teacher training institutes involved in the study. This suggests that student-teachers' disposition toward realistic modeling of arithmetic

word problems is at least partially influenced by the courses on mathematics education received during their pre-service training.

Test 2. The student-teachers' strong disposition toward non-realistic modeling was also revealed by their evaluations of the realistic answer (RA) and the non-realistic answer (NA) on the same P-items during Test 2, as is shown in Table 5.

Insert Table 5 here

A t-test revealed that the difference between the percentage of 1-scores for the NA (56%) and for the RA (47%) was significant (t-test, two-tailed, $t=6.25$, $p<.01$); the same holds for the difference between the percentage of 0-scores for the NA (18%) and for the RA (47%) (t-test, two-tailed, $t=-21.56$, $p<.01$). Thus, the student-teachers overall evaluation of the stereotyped, non-realistic answer to the P-items was considerably more positive than for the realistic answer based on context-based considerations.

There was again a significant difference between the first-year and the third-year student-teachers. The third-year students gave significantly more 1-scores (t-test, two-tailed, $t = 3.29$, $p < .01$) and less 0-scores (t-test, two-tailed, $t = 2.63$, $p < .01$) for the RAs than the first-year students. Correspondingly, the third-year students produced significantly less 1-scores (t-test, two-tailed, $t = 2.30$, $p < .05$) and more 0-scores (t-test, $t = 2.33$, $p < .05$) for the NAs than the first-year students. As for Test 1, the size of the difference between the first-year and the third-year students was not the same in the three teacher-training institutes.

Relationship between Test 1 and Test 2. The previous sections focused on the results for Test 1 and Test 2 separately. However, it is also interesting to examine to what extent the student-teachers' evaluations of the NAs and RAs during Test 2 matched their own performances on Test 1. This was done by analyzing the scores for the RA and the NA following the 52% non-realistic reactions (NRs) and the 48% realistic reactions (RRs) on Test 1 separately (see Table 6).

Insert Table 6 here

The left part of Table 6 presents the distribution of the different combinations of RA scorings (1, 1/2 or 0) and NA scorings (1, 1/2 or 0) over the seven P-items of Test 2 for the 52 % non-realistic reactions on Test 1, as well as the distribution of the scorings for the RA and the NA over the three scores (1, 1/2 and 0).

As expected, we found a strong relationship between the non-realistic reactions on a P-item during Test 1, and the evaluations of the RA and the NA on that item during Test 2. In 89.3 % of all cases wherein a NR was given to a P-item during Test 1, the NA to that item was given a 1-score in Test 2. Correspondingly, 83.1 % of the NRs during Test 1 were followed by a 0-score for the RA in Test 2. The match of a 1-score for the NA and a 0-score for the RA occurred in no less than 79.3 % of all cases wherein a P-item from Test 1 was answered with a NR. Apparently, the NA was scored with 1 because this response corresponded to the stereotyped, non-realistic answer the student-teachers had given themselves on this item during Test 1, and they scored the RA with 0 because they could not understand, and, therefore, appreciate the context-based considerations underlying this answer.

Ten percent of the NRs to a P-item during Test 1 were followed by a 1-score for the RA during Test 2. This suggests that in those cases the confrontation with the RA during Test 2 had functioned as a scaffold toward (more) realistic modeling. However, the finding that only 10.0 % of the scorings following a NR yielded evidence for the scaffolding effect of the confrontation with the RA, can be interpreted as additional evidence for the strength and the resistance of the tendency toward non-realistic mathematical modeling among student-teachers.

The right part of Table 6 presents the distribution of the different combinations of RA scorings (1, 1/2 or 0) and NA scorings (1, 1/2 or 0) over the seven P-items of Test 2 following the 48 % RRs on Test 1, as well as the distribution of the scorings for the RA and the NA over the three scores (1, 1/2 and 0).

As shown in the Table, the congruence between the RRs on Test 1 and the scorings of the RA and the NA during Test 2 was less straightforward than for the NRs. On the one hand, the evaluations of the RA were generally in line with the reactions on Test 1: indeed, 85.4 % of the RRs on Test 1 were followed by a 1-score for the RA on Test 2.

But the scorings for the NA were rather surprising: only 33.5 % of the subjects who reacted in a realistic way to a P-item during Test 1 scored the NA with a 0 during Test 2 (almost always in combination with a 1 for the RA). This indicates that in many cases where student-teachers reacted themselves to a P-item in a realistic manner, they were nevertheless quite understanding and tolerant to elementary school pupils who interpreted and solved these P-items without seriously taking into account the relevant knowledge about the context called up by the problem

statement. According to their written explanations in the comments box, they thought that it is unfair to punish a fifth-grader for solving the P-item in a stereotyped, non-realistic manner. This is illustrated by the following comment accompanying the scoring combination "1 for RA and 1 for NA" with respect to the runner-item: "I scored alternative D (= the RA: "It is impossible to answer precisely what John's best time on 1 kilometer will be") with 1 because this pupil who gave this answer knows that is not realistic to assume that John will be able to run at his record speed for 1 kilometer. But I also gave 1 for alternative A (= the NA: " $17 \times 10 = 170$. John's best time to run 1 kilometer is 170 seconds") because from a purely computational point of view this is the correct answer."

Discussion

This study demonstrates convincingly that pre-service teachers are themselves strongly affected by the phenomenon of unrealistic modeling in word problem solving. Of course, this finding does not provide direct evidence that these teacher conceptions and beliefs are responsible for pupils' strong tendency to exclude real-world knowledge from their solution processes of arithmetic word problems. However, taking into account recent research on mathematics teaching and learning (see e.g., De Corte et al, 1996), it is plausible to assume that these teacher cognitions and beliefs concerning the role and the importance of real-world knowledge in interpreting and solving verbal problems, have a substantial impact on their actual teaching behavior, and, consequently, on their pupils' learning activities and outcomes. Therefore, if one wants to link arithmetic word problem solving to the everyday experiential world of the children - as is strongly advocated in current reform documents relating to mathematics education - it will be necessary to stimulate and support (future) teachers to develop the appropriate conceptions and beliefs that are needed in view of (re-)orienting their teaching practices toward realistic modeling of problem situations and toward realistic interpretations of the outcomes of arithmetic operations, as components of a genuine mathematical disposition.

An intervention study: Teaching realistic modeling to fifth graders

Starting from the preceding studies about the phenomenon of unrealistic modeling, and taking into account the literature on mathematics learning and teaching (De Corte et al., 1996), an intervention study was designed. Using a pretest-posttest design with two control groups a teaching experiment was carried out to verify the hypothesis, that through appropriate instruction in a powerful learning environment it is possible to develop in

primary school children a disposition toward realistic modeling of arithmetic word problems (for a more detailed report, see Verschaffel & De Corte, in press). In other words, we assumed that pupils can learn good things from better instruction.

Method

Three classes from the same school participated in the experiment: one experimental (E) of 19 fifth-grade children, and two control classes (C1 and C2) of 18 and 17 sixth-grade children, respectively. The pupils from the E-class were taught an experimental program on realistic modeling (during the regular hours allocated for mathematics teaching). The program consisted of five teaching-learning units (TLU) of 2 1/2 hours each. During the experiment the pupils from the two control classes followed the regular mathematics curriculum.

The three classes were administered a pretest consisting of 10 items in which the mathematical assumptions were problematic (P-items), and five standard problems (S-items) as buffer items. With respect to each problem, pupils were not only asked to write their response, but also to mention how they arrived at their answer and/or possible difficulties or worries experienced during the solution of the problem. At the end of the experimental course in the E-class, a parallel version of the pretest was administered as a posttest in the three groups. However, in one of the control classes (C1) this posttest was preceded by an introduction of 15 minutes in which the pupils' attention was drawn to the fact that routine solutions for word problems are sometimes inappropriate when considered in terms of realistic constraints. A few examples of such inappropriate routine solutions were given, and pupils were warned that the test contained several items for which such routine solutions are inappropriate. One month after the posttest the pupils of the E-class were given a retention test, involving again 10 P-problems.

Analysis of variance was used to evaluate the effects of the experimental program on children's realistic reactions on the P-problems. As in the previous studies a response was considered as realistic, either because a realistic answer was given, or a non-realistic answer was accompanied by a realistic comment.

Major features of the powerful learning environment

The experimental course was developed in line with our design principles for powerful teaching-learning environments (see De Corte, 1995b; De Corte et al, 1996).

First, the impoverished and stereotyped diet of standard word problems offered in traditional mathematics classrooms, was replaced by a set of more authentic problem situations especially designed to stimulate pupils to pay attention at the complexities involved in realistic mathematical modeling, and at distinguishing between realistic and stereotyped solutions of mathematical applications. Each TLU focused on one prototypical problematic topic of realistic modeling. The topic of the first TLU was: making appropriate use of real-world knowledge and realistic considerations when interpreting the outcome of a division problem involving a remainder. The opening problem involved a story about a regiment of 300 soldiers doing several military activities. Each part of the story was accompanied with a question which always asked for the same arithmetic operation (namely, $300 : 8 = .$) but required each time a different answer (respectively, "38", "37", "37.5" and "37 remainder 4"). The theme of the second TLU was the union or separation of two sets with joint elements. The opening problem was an (unsolvable) problem about a boy who possessed already a given number of comic strips from the popular series "Suske and Wiske" and got a package of second-hand albums of "Suske and Wiske" from his older cousins (who had lost their interest in these comics). The pupils had to determine how many albums of "Suske and Wiske" there were still missing in the boy's collection after getting this present, given the total number of available albums in this series (i.e., 236), the number of albums in the boy's collection before the gift (i.e., 96), and the number of albums in the package he got from his cousins (i.e., 45). The third TLU focused on problem situations wherein it is not immediately clear whether the result of adding or subtracting two given numbers yields the appropriate answer or an answer that is 1 more or 1 less than the correct one. In the opening problem, pupils were given the number of the first and the last ticket sold at the cash desk of a swimming pool on a particular day (i.e., ticket no. 524 and no. 616, respectively), and they had to decide how many tickets were sold that day. The fourth TLU started with a problem about a boy who wanted to make a swing and had to decide about the amount of rope needed for fastening the swing at a horizontal branch of a big old tree at a height of 5 meter. In that session, pupils had to experience that in many application problems one has to take into account several relevant elements that are not explicitly nor immediately "given" in the problem statement but that belong to one's common-sense knowledge base. The fifth and last TLU dealt with the principle of proportionality and, more particularly, with how to discriminate between cases where solutions based on direct proportional reasoning are or are not appropriate.

The starting problem was about a young athlete whose best time to run 100 meters was given, and pupils were asked to predict his best time on the 400 meters.

Second, not only the problems but also the teaching methods differed considerably from traditional mathematics classroom practice. The opening problem of each TLU was solved in mixed-ability groups of 3-4 pupils. This group assignment was followed by a whole-class discussion, in which the answers, the solution processes and possible additional comments of the different groups were compared. Then, each mixed-ability group was given a set of four or five new problems, some with and some without the same underlying modeling difficulty as the opening problem. This group assignment was again followed by a whole-class discussion. Finally, each pupil was individually administered one problem that involved once again the topical modeling difficulty. Pupils' reactions to this individual assignment were also discussed afterwards during a whole-class discussion.

Third, a new classroom culture was created featuring systematic attention to the realistic modeling perspective as an aspect of acquiring a mathematical disposition. This was attempted by activities such as modeling, explaining, and discussing valuable problem-solving strategies, and listening carefully to children's explanations and justifications of their own solution procedures. Another major dimension of establishing a new classroom culture consisted of explicitly negotiating with the pupils new socio-mathematical norms about the role of the teacher and the children in the classroom, and about what counts as a good mathematical problem, a good solution procedure, a good response (see also Cobb, Yackel, & Wood, 1992; Gravemeijer, 1995).

Hypotheses

First, in line with the results of Study 1 and 2, we hypothesized that on the pretest the pupils would demonstrate a strong tendency to exclude real-world knowledge when confronted with the problematic versions of the problems. Therefore, we predicted that on the P-items of the pretest the pupils of the three classes (E, C1 and C2) would produce a percentage of RRs that would not differ significantly from the 17 % RRs and 16 % RRs obtained in the first and the second studies, respectively.

Second, we hypothesized that the experimental program would induce conceptual change in students resulting in a disposition toward realistic modeling and interpreting of arithmetic word problems. Therefore, we predicted a significant increase in the number of RRs on the P-items from pretest to posttest for the E-group. For the two other groups no significant increase from pretest to posttest was expected.

A third hypothesis was that the positive effect of the experimental program would be lasting. Consequently, we predicted that there would be no significant difference between the percentage of RRs of the E-group on the P-items of the retention test and the parallel problems of the posttest.

Results

As predicted in the first hypothesis, the pupils from all three classes demonstrated a strong overall tendency to exclude real-world knowledge from their problem solutions during the pretest. Overall, the percentage of RRs on the 10 P-items of the pretest was only 15 %. The fifth-graders from the E-class produced less RRs than the sixth-graders from the C1- and the C2-classes: the percentages were 7 %, 20 % and 18 %, respectively. But according to a Tukey a posteriori tests the differences were not significant.

To test the second hypothesis, we analyzed whether there was an interaction effect between the independent variables "Group" (E, C1 and C2) and "Time" (pretest versus posttest). The analysis of variance revealed that this "Group x Time" interaction was indeed significant ($p < .0001$). A Tukey a posteriori tests showed that there was a significant increase in the number of RRs from pretest to posttest for the E-group: from 7 % RRs on the pretest to 51 % RRs on the posttest. To the contrary, in the two control classes the progress in the number of RRs from pretest to posttest was non-significant, namely from 20 % to 34 % for C1, and from 18 % to 23 % for C2 (see Table 7).

Insert Table 7 here

Finally, the positive results of the E-class on the retention test confirmed the third hypothesis. The percentage of RRs for the items of the retention test that were parallel with the posttest was 40 %, which is almost the same as the result for those items on the immediate posttest (i.e., 41 %).

General discussion

In this presentation five related studies on upper elementary school children's conceptions about the role of real-world knowledge in mathematical modeling of school word problems, were reported.

The findings from the first study show that - as a consequence of current classroom practice in teaching arithmetic word problem solving - pupils have gradually built up a very strong tendency to solve arithmetic word problems in a stereotyped way: they apply one (or a combination) of the basic arithmetic operations with the two given numbers in the problem, without any consideration of the possible problematic modeling assumptions underlying their proposed solution. The results of the second and the third study confirm this tendency among upper elementary school children, and show that it is culturally independent. Those studies also indicate that the tendency is rooted in very strong conceptions and beliefs in pupils about the role of real-world knowledge in mathematical modeling of school word problems, which are resistant to conceptual change. However, the investigations revealed likewise considerable interindividual differences in pupils' actual level of performance with respect to the use of everyday knowledge in mathematical modeling (see Study 1), as well as in their learning potential (or their zone of proximal development) as reflected in their reactions to simple forms of scaffolding aimed at triggering the activation of that knowledge (see Study 2).

The fourth study revealed that upper elementary school pupils' strong tendency to solve arithmetic word problems in a stereotyped way was paralleled by a similar tendency among (future) teachers, as reflected by their own preferred spontaneous solutions to word problems and their evaluations of different types of pupil answers. It seems reasonable that these teacher conceptions and beliefs are at least partially responsible for the tendency toward non-realistic modeling and problem solving among pupils.

The fifth study assessed whether it is possible to change children's inappropriate conceptions and beliefs about mathematical modeling and about the role of real-world knowledge in it. In view of inducing change in their conceptions and beliefs, a group of pupils was immersed into a new classroom culture in which word problems were effectively treated as exercises in realistic mathematical modeling. In this new classroom culture a set of challenging, non-standard problem situations and a number of instructional techniques like small-group work and whole-class discussions were systematically used to produce a cognitive conflict (or disequilibrium) in the pupils, which was considered necessary to produce the desired revision of their beliefs about the role of everyday knowledge in mathematical applications. The findings provide good and promising support for the hypothesis that it is possible to change pupils' conceptions about and their disposition toward the activation of real-world knowledge in mathematical modeling of school word problems.

References

- Burkhardt, H. (1994). Mathematical applications in school curriculum. In T. Husén & T.N. Postlethwaite (Eds), The international encyclopedia of education (2nd ed.) (pp. 3621-3624). Oxford/New York: Pergamon Press.
- Cai, J., & Silver, E.A. (1994). Solution processes and interpretations of solutions in solving a division-with-remainder story problem: Do Chinese and U.S. students have similar difficulties? Journal for Research in Mathematics Education, 26, 491-497.
- Cobb, P., Yackel, E., & Wood, T. (1992). A constructivist alternative to the representational view of mind in mathematics education. Journal for Research in Mathematics Education, 23(1), 2-33.
- De Corte, E. (1995a). Designing powerful teaching-learning environments conducive to the acquisition of cognitive skills. In R. Olechowski & G. Khan-Svik (Eds.), Experimental research on teaching and learning (pp. 67-82). Frankfurt am Main: Peter Lang.
- De Corte, E. (1995b). Fostering cognitive growth: A perspective from research on mathematics learning and instruction. Educational Psychologist, 30, 37-46.
- De Corte, E., Greer, B., & Verschaffel, L. (1996). Psychology of mathematics teaching and learning. In D.C. Berliner & R. Calfee (Eds.), Handbook of educational psychology. New York: MacMillan.
- De Corte, E., & Verschaffel, L. (1985). Beginning first graders' initial representation of arithmetic word problems. Journal of Mathematical Behavior, 4, 3-21.
- De Corte, E., & Verschaffel, L. (1989). Teaching word problems in the primary school. What research has to say to the teacher. In B. Greer & G. Mulhern (Eds), New developments in teaching mathematics (pp. 85-106). London: Routledge.
- Freudenthal, H. (1991). Revisiting mathematics education. Dordrecht, The Netherlands: Kluwer.
- Gravemeijer, K.P.E. (1995). Developing realistic mathematics education. Utrecht: Freudenthal Institute, University of Utrecht.
- Greer, B. (1993). The modeling perspective on word problems. Journal of Mathematical Behavior, 12, 239-250.
- Hamers, J.H.M., Sijtsma, K., & Ruijsenaars, A.J.J.M. (1993). Learning potential assessment: Theoretical, methodological and practical issues. Amsterdam/Lisse: Swets & Zeitlinger.
- Kilpatrick, J. (1987). Problem formulating: Where do good problems come from? In A.H. Schoenfeld (Ed.), Cognitive science and mathematics education (pp. 123-147). Hillsdale, NJ: Erlbaum.
- McLeod, D.B., & Adams, V.M. (1989). Affect and mathematical problem solving. New York, NY: Springer-Verlag.
- Reusser, K. (1988). Problem solving beyond the logic of things: Contextual effects on understanding and solving word problems. Instructional Science, 17, 309-338.

- Reusser, K. (in press). The suspension of reality and sense-making in the culture of school mathematics: The case of word problems. Learning and Instruction.
- Robitaille, D.F., & Travers, K.J. (1992). International studies of achievement in mathematics. In D.A. Grouws (Ed.), Handbook of research on mathematics teaching and learning (pp. 687-709). New York: Macmillan.
- Säljö, R. (1991). Learning and mediation. Fitting reality into a table. Learning and Instruction, 1, 261-273.
- Schoenfeld, A.H. (1988). When good teaching leads to bad results: The disasters of "well-taught" mathematics courses. Educational Psychologist, 23, 145-166.
- Selter, C. (1994). How old is the captain? Strategies, 5(1), 34-37.
- Silver, E.A., Shapiro, L.J., & Deutsch, A. (1993). Sense making and the solution of division problems involving remainders: An examination of middle school students' solution processes and their interpretations of solutions. Journal for Research in Mathematics Education, 24, 117-135.
- Stevenson, H.W., & Stigler, J.W. (1992). The learning gap. Why our schools are failing and what can we learn from Japanese and Chinese education. New York: Summit Books.
- Treffers, A. (1987). Three dimensions. A model of goals and theory description in mathematics education. The Wiskobas project. Dordrecht, The Netherlands: Reidel.
- Verschaffel, L., & De Corte, E. (in press). Teaching realistic mathematical modeling in the elementary school. A teaching experiment with fifth graders. Journal of Research in Mathematics Education.
- Verschaffel, L., De Corte, E., & Borghart, I. (in press). Pre-service teachers' conceptions and beliefs about the role of real-world knowledge in mathematical modeling of school word problems. Learning and Instruction.
- Verschaffel, L., De Corte, E., & Lasure, S. (1994). Realistic considerations in mathematical modeling of school arithmetic word problems. Learning and Instruction, 4, 273-294.
- Vygotsky, L.S. (1978). Mind in society. The development of higher psychological processes. Cambridge, MA : Harvard University Press.

Table 1. Ten P-items involved in Study 1

P1 Carl has 5 friends and Georges has 6 friends. Carl and Georges decide to give a party together. They invite all their friends. All friends are present. How many friends are there at the party? (= the "birthday" item)

P2 Steve has bought 4 planks of 2.5 meter each. How many planks of 1 meter can he get out of these planks? (= the "planks" item)

P3 What will be the temperature of water in a container if you pour 1 liter of water at 80° and 1 liter of water of 40° into it? (= the "water" item)

P4 450 soldiers must be bused to the their training site. Each army bus can hold 36 soldiers. How many buses are needed? (= the "buses" item)

P5 John's best time to run 100 meters is 17 seconds. How long will it take to travel 1 kilometer? (= the "runner" item)

P6 Bruce and Alice go to the same school. Bruce lives at a distance of 17 kilometers from the school and Alice at 8 kilometers. How far do Bruce and Alice live from each other? (= the "school distance" item)

P7 Grandfather gives his 4 grandchildren a box containing 18 balloons, which they share equally. How many balloons does each grandchild get? (= the "balloons" item)

P8 Rob was born in 1978. Now it's 1993. How old is he? (= the "age" item)

P9 A man wants to have a rope long enough to stretch between two poles 12 meters apart, but he has only pieces of rope 1.5 meters long. How many of these pieces would he need to tie together to stretch between the poles? (= the "rope" item)

P10 This flask is being filled from a tap at a constant rate. If the depth of the water is 4 cm after 10 seconds, how deep will it be after 30 seconds? (This problem was accompanied by a drawing of a cone-shaped flask) (= the "flask" item)

Table 2. Number and percentage of non-realistic (NR) and realistic (RR) pupil reactions to the ten P-items in Study 1

P-items	NR		RR	
	N	%	N	%
1	60	80	15	20
2	65	87	10	13
3	62	83	13	17
4	38	51	37	49
5	73	97	2	3
6	73	97	2	3
7	31	41	44	59
8	73	97	2	3
9	75	100	0	0
10	72	96	3	4
Total	622	83	128	17

Table 3. Frequency and cumulative frequency, and percentage and cumulative percentage of realistic reactions of the 30 pupils to the seven problematic items at the different stages of Study 2

		IRR*	RRF	RRS	NRR
Absolute	N	49	34	37	90
	%	23%	16%	18%	43%
Cumulative	N	49	83	120	210
	%	23%	40%	57%	100%

- * IRR = immediate realistic reaction during the collective test
 RRF = realistic reaction after the first, weak scaffold
 RRS = realistic reaction after the second, strong scaffold
 NRR = still a non-realistic reaction at the end of the individual interview

Table 4. Number (N) and percentage (%) of RR to the ten P-items in the three groups

		P-items																					
		1		2		3		4		5		6		7		8		9		10		Total	
		N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Japan																							
Condition I	(N=45)	16	13	0	0	5	11	28	62	3	7	1	2	23	52	0	0	1	2	3	4	80	15
Condition II	(N=46)	7	15	4	9	8	17	28	61	9	20	3	7	35	76	2	4	0	0	1	2	97	20
Belgium																							
(N=75)		15	20	10	13	13	17	37	49	2	3	2	3	44	59	2	3	0	0	3	4	128	17

Table 5. Student teachers' evaluations of pupils' realistic (RA) and non-realistic answers (NA)

Year	RA			NA		
	1	1/2	0	1	1/2	0
1st	44	7	49	58	25	17
3rd	53	5	42	52	26	22
1st + 3rd	47	6	47	56	26	18

Table 6. Combinations of RA scorings (1, 1/2 or 0) and NA scorings (1, 1/2 or 0) over the seven P-items of Test 2 for the total number of non-realistic reactions (52 %) and of realistic reactions (48 %) on Test 1

NON-REALISTIC REACTIONS					REALISTIC REACTIONS				
NA	RA				NA	RA			
	1	1/2	0	Total		1	1/2	0	Total
1	3.4	6.6	79.3	89.3	1	10.2	3.2	7.4	20.8
1/2	3.8	0.1	2.6	6.5	1/2	42.3	2.0	1.4	45.7
0	2.8	0.2	1.2	4.2	0	32.9	0.4	0.2	33.5
Total	10.0	6.9	83.1	100.0	Total	85.4	5.6	9.0	100.0

Table 7. Number and percentage of realistic reactions (RRs) on the 10 P-items from the pretest and the posttest in the experimental (E) class and the two control classes (C1 and C2)

Groups	Realistic reactions (RRs)			
	<u>Pretest</u>		<u>Posttest</u>	
	N	%	N	%
E (n=18*)	13	7 %	91	51 %
C1 (n=18)	36	20 %	61	34 %
C2 (n=17)	32	18 %	39	23 %

* Due to illness one pupil from the E-group could not participate in the posttest.

邦子さんは、それぞれ2.5 mの長さの板を4つ買いました。
これらの板から1 mの板がいくつとれるでしょう？。

Figure 1. Example of the P2 problem (Kuniko has bought 4 planks of 2.5 m each. How many planks of 1 m can she saw out of these planks?) as it was presented in the study.

This flask is being filled from a tap at a constant rate. If the depth of the water is 4 cm after 10 seconds, how deep will it be after 30 seconds? (This problem was accompanied by a figure of a partly filled cone-shaped flask)

$$3 \times 3.5 \text{ cm.} = 11.5 \text{ cm.}$$

After 30 seconds the depth of
the water will be 11.5 cm.

A

$$3 \times 3.5 \text{ cm.} = 10.5 \text{ cm.}$$

After 30 seconds the depth
of the water will be 10.5 cm.

B

$$3.5 \text{ cm.} + 20 \text{ cm.} = 23.5 \text{ cm.}$$

After 30 seconds the depth of
the water will be 23.5 cm.

C

It is impossible to give a
precise answer.

D

Figure 2. Presentation of problems in Test 2 of Study 3